◆ Human Brain Mapping 38:5706–5725 (2017) ◆

Taking Tests in the Magnet: Brain Mapping Standardized Tests

David C. Rubin,[1,2]* **Dawei Li**,[1] **Shana A. Hall**,[1] **Philip A. Kragel**,[3] **and Dorthe Berntsen**[2]

[1]*Department of Psychology & Neuroscience, Duke University, Durham, North Carolina*
[2]*Center on Autobiographical Memory Research, Aarhus University, Denmark*
[3]*Institute of Cognitive Science, University of Colorado Boulder, Boulder, Colorado*

Abstract: Standardized psychometric tests are sophisticated, well-developed, and consequential instruments; test outcomes are taken as facts about people that impact their lives in important ways. As part of an initial demonstration that human brain mapping techniques can add converging neural-level evidence to understanding standardized tests, our participants completed items from standardized tests during an fMRI scan. We compared tests for diagnosing posttraumatic stress disorder (PTSD) and the correlated measures of Neuroticism, Attachment, and Centrality of Event to a general-knowledge baseline test. Twenty-three trauma-exposed participants answered 20 items for each of our five tests in each of the three runs for a total of 60 items per test. The tests engaged different neural processes; which test a participant was taking was accurately predicted from other participants' brain activity. The novelty of the application precluded specific anatomical predictions; however, the interpretation of activated regions using meta-analyses produced encouraging results. For instance, items on the Attachment test engaged regions shown to be more active for tasks involving judgments of others than judgments of the self. The results are an initial demonstration of a theoretically and practically important test-taking neuroimaging paradigm and suggest specific neural processes in answering PTSD-related tests. *Hum Brain Mapp 38:5706–5725, 2017.*

Key words: cognition; psychological theory; psychology; applied; Theory of Mind; anxiety; personality

INTRODUCTION

Standardized psychometric tests (STs) assume that people vary in fairly stable ways that can be assessed in a brief period of time on measures of their knowledge, abilities, traits, opinions, and health. Great effort and statistical sophistication go into the construction of these tests; they are often either the operational definition of the concept they measure or the best brief behavioral measure of it. The results of such tests are more than numbers. They are taken as facts about people that determine what schools they can attend, what jobs they can obtain, and what diagnoses they might be given. Among the standard tools used to better understand the constructs that tests measure are statistical analyses of test items, epidemiological studies of how test scores vary over populations, behavioral manipulations of the conditions under which tests are taken, and correlations of the test scores with behavior, neural structure, and neural function.

A neglected source of information about STs is the neuroimaging of the neural systems active when such tests are being taken. Such data could be especially useful

Contract grant sponsor: NIMH; Contract grant number: R01-MH066079; Contract grant sponsor: Danish National Research Foundation; Contract grant number: DNRF89

*Correspondence to: David C. Rubin, Psychology and Neuroscience, Duke University, Durham, NC 27708-0086, USA. E-mail: david.rubin@duke.edu

Received for publication 3 January 2017; Revised 10 July 2017; Accepted 28 July 2017.

DOI: 10.1002/hbm.23761

Published online 18 August 2017 in Wiley Online Library (wileyonlinelibrary.com).

because we already have a good theoretical understanding of the neural localization of many of the cognitive and emotional processes that are relevant to the taking of tests. If a literature of neural studies of STs emerged, more sophisticated comparisons directly related to specific tests and the conditions under which they are taken would become available. Between- and within-subject manipulations are also possible. For instance, the effects on neural processes of the stereotype manipulation of informing older adults that an ST of a cognitive ability was a measure of memory versus a measure of wisdom could be examined to see which neural processes the manipulation affected and which varied parametrically with performance on the ST. The existing knowledge from both the behavioral and neural levels make STs an ideal topic to integrate across these levels of analysis [Schwartz et al., 2016].

Because of the relevance of PTSD to many areas and our previous contributions to it [Rubin et al., 2008, 2014; St. Jacques et al., 2011, 2013], we chose four extensively studied classes of STs that measure or are correlated with PTSD, including the Posttraumatic Stress Disorder Check List (PCL-5) [Weathers et al., 2013], the Centrality of Events Scale (CES) [Berntsen and Rubin, 2006, 2007], Neuroticism [Costa and McCrae, 2008; John et al., 1991], and Attachment [Mikulincer and Shaver, 2007]. The PCL-5 is one of the most commonly used screening tests for PTSD. Each of its 20 items closely follows the 20 symptoms of the diagnosis [American Psychiatric Association, 2013]. When a licensed clinician determines that a minimum number of these 20 symptoms related to a negative event is present in each of four categories, that they have serious effects on normal functioning, and that the event is a trauma, a PTSD diagnosis is given. The categories include intrusion symptoms related to recurrent intrusive memories or dreams and negative reactions to them, avoidance symptoms involving effort to prevent such memories, negative alterations in cognition or mood, and increased arousal. The CES measures the extent to which individuals construe a stressful or traumatic event as a central component of their narrative identity and sense of self. Items measure the extent to which a traumatic event colors how other events are viewed, signals a major turning point in the life story, and has become a core component of identity. The CES has strong empirical support in correlating with PTSD and symptoms of posttraumatic stress across a variety of participant samples and trauma types, even after controlling for other known risk factors of PTSD symptoms, such as anxiety, depression, and dissociation [Berntsen and Rubin, 2007; Rubin et al., 2014].

Unlike the PCL and CES, Neuroticism and Attachment are not about reactions to single events. Of the small set of traits or dimensions commonly used to describe personality with STs [Costa and McCrae, 2008; John et al., 1991], Neuroticism is the trait most often associated with psychological disorders in general and shows consistent correlations with PTSD [Rubin et al., 2014]. For instance, a review of 11 studies totaling 1,415 individuals found Neuroticism had an average correlation of 0.43 with PTSD symptom severity [Rubin et al., 2008]. Attachment refers to systematic patterns of expectations, beliefs, and emotions about the availability and responsiveness of close others during times of distress. Attachment originated from theories of development rather than personality [Bowlby, 1969; Mikulincer and Shaver, 2007]. Insecure attachment can include discomfort with close interpersonal relationships, persistent fear of interpersonal rejection or abandonment, and excessive concerns about and desire for proximity to close others. It can increase vulnerability to PTSD symptoms (Charuvastra and Cloitre [2008] and Ogle et al. [2015] for a review).

To control for the general processes involved in test taking, including reading, comprehension, decision-making, and responding, we included a test of General Knowledge. Participants agree or disagree with items such as "Sydney is the capital of Australia." We chose this task because it contains general information that does not depend on the context in which it was learned and is not about the participant. In addition, the items have little relevance to the psychopathology, personality, or the participant's general way of behaving. In terms of cognitive theory, it provides a semantic memory task to contrast with the STs, which are more episodic.

We could not find systematic attempts to measure STs during neuroimaging and therefore could not justify a priori anatomically specific hypotheses. There are studies analyzing how neural activity exhibited by individuals, who vary on their scores on STs, varies across different tasks. For STs such as the ones used here, these include tasks related to resting state [Brown et al., 2014], autobiographical memory, and PTSD. These studies will be useful in forming hypotheses in future work once we can establish some basic results here. However, these tasks bear no direct relation to the neural processes involved in taking the STs, and so using them to form hypotheses about the neural basis of test taking is hard to justify. We therefore adopt the strategy of reporting our results with minimal speculation, interpreting neural activity using meta-analyses of neural activity related to particular areas or tasks rather than individual studies. That is, we use statistically rigorous descriptions followed by theoretically and empirically informed interpretations, while trying to avoid reverse inferences.

We begin with a pattern classification analysis testing whether differences in neural activity can be used to predict which test a participant is taking based on the activity of other participants. This analysis requires no a priori knowledge of the areas involved. Thus, we quantify the extent to which the STs engage different neural processes before proceeding with the more detailed analyses of activation within individual areas. Together, these analyses aim to characterize unique and shared processes engaged when taking STs.

METHODS

Participants

At the final session for an fMRI study of involuntary memory in PTSD, we recruited all participants who provided data to obtain participants who were familiar and comfortable with the fMRI environment. The study from which we recruited is similar to one we recently published [Hall et al., 2014]. In it participants learned pairings of environmental sound and pictures; in the scan they heard the sound and decided if it came from left or right of midline, and after the scan they reported whether picture came to mind during the scan. The new study followed the same involuntary memory procedure except that it included participants with PTSD and negative and neutral pictures. The pictures were not selected to be relevant to the participants' traumas or to their intrusive memories.

Twenty-seven participants agreed to take part in this additional study. They gave written informed consent for a protocol approved by the Duke University Medical Center Institutional Review Board. All participants were trauma-exposed, right-handed, fluent English speakers, with normal or corrected-to-normal vision. We limited participants to those who were trauma exposed because we were asking about reactions to traumas; however, this should not make our sample very different from the general adult USA population, which is 85% trauma exposed [Kilpatrick et al., 2013; Roberts et al., 2011]. Participants were excluded if they had any contraindications to MRI (e.g., metallic implants), a head injury with loss of consciousness, any psychiatric diagnoses other than PTSD, or were currently taking psychotropic medication. The participants were all screened for PTSD by a staff member trained at the Veteran's Administration Hospital to administer the Clinician Administered PTSD Scale (CAPS) [Blake et al., 1995], a standard test for determining PTSD, for clinical and research purposes. The index traumas for the CAPS for the participants with and without a resulting PTSD diagnosis were: injury, illness, or accident (4, 4), sexual assault (3, 4), injury or death to a family member or friend (2, 4), physical assault (4, 0), witnessed others death (0, 1), and abortion (0, 1).

Four participants were excluded due to excessive motion. Our criterion was any scan with >5 mm of motion or participant with more than 5% of their total scans with between 3 and 5 mm, though the four excluded participants actually had more than 10% of their total scans >5 mm of motion or more than 30% of their total scans >3 mm of motion. The remaining participants had a mean age of 21.74 (SD = 3.32, range 18–31). Twelve were male, 13 were White, 5 were Black, and 5 were Asian. Their scores on the CAPS were (M = 31.74, SD = 21.74, range 0–79); 11 met the criteria for current PTSD.

Materials

To adapt the STs to our study, we formed modified Standardized Tests (mSTs) in the following manner. We kept items and response scales on all five questionnaires as close as possible to the original sources, though more liberties were taken with the General Knowledge questions, which had not been used as a standardized individual differences test. All response scales were reduced to a four-point rating scale to match the four response keys on the button box used in the scanner. Most scales were initially five-point scales and for these we removed the middle category. This left the PCL with the four response options: not at all, a little bit, quite a bit, and extremely. For the CES, we added labels of disagree and agree to the existing extreme values of totally disagree and totally agree.

We included 20 items for each mST in each of three runs in the scanner. Because there were only 60 items for each mST, we did not attempt to divide the mSTs into subscales. The PCL and CES each have 20 items, so in each run participants responded to a different stressful or traumatic event. Participants had nominated three negative events during earlier testing, described in the beginning of the Participant section. The order of presentation of these events was randomized for each participant.

For Attachment and Neuroticism, we selected 60 items from multiple STs because no one standardized test had enough items. For these, the items from each standardized test were given in the order they appeared in the standardized test, but items from the various tests were interspersed so that each run had an approximately equal number of items from each test. The 60 Neuroticism items included all 48 Neuroticism items from the NEO Personality Inventory-Revised (NEO-PI-R) [Costa and McCrae, 2008], all 8 Neuroticism items from the Big-5 Inventory [John et al., 1991], and the 4 items that had the least direct overlap with the other two tests from the Generalized Anxiety Disorder 7-item (GAD-7) scale [Spitzer et al., 2006]. Items from the latter two tests were changed to begin with statements such as "I am often" or "I often feel" to match the NEO items in form. Participants responded on a scale of disagree strongly, disagree a little, agree a little, and agree strongly, modified from the NEO. The 60 Attachment items included all 36 items of the Experiences in Close Relationships Inventory (ECR) [Brennan et al., 1998] and the 24 of the 30 items of the Relationship Styles Questionnaire (RSQ) [Griffin and Bartholomew, 1994] that had the least direct overlap. Participants responded on a scale of strongly disagree, slightly disagree, slightly agree, and strongly agree, which were scale items 1, 3, 5, and 7 on the 7-point scale used by the ECR.

We devised the 60 General Knowledge items from existing norms from a college student population [Tauber et al., 2013]. The response scale was wrong, probably wrong, probably right, and right, with approximately half of the items changed to be incorrect using common errors published in the norms. As a way of increasing attention to the task, we used a pseudorandom order for the items; several obviously incorrect items occurred early in the first run to alert the participants that not all items would be correct.

In each of the three runs, we used the same order of General Knowledge, Attachment, Neuroticism, CES, and PCL to stress to the participants that the same five types of questions were being asked repeatedly. We started with the control task which asked about the widest range of information, then went to the two mSTs that asked about the participants' general tendencies in a wide range of situations, before the two mSTs that asked about specific stressful events. We chose this order to minimize the chances of the specific events affecting responses to questions about more general behavioral tendencies.

Procedure

Before entering the scanner, we asked participants to think of the three most negative events they had experienced. They gave these events a short name and wrote a paragraph to describe the event. The names were used later to cue the memory about which the participants answered the PCL and the CES. We informed the participants that they would be asked to respond to five types of questions and that these would repeat three times. We then gave instructions on the button box scales, provided instructions for and examples of each question type, and explained that for two scales, the event given would be selected from one of the negative experiences that they listed previously. Participants were also informed that each item would appear by itself for 5 s, then with a scale for 1.5 s, and that they should try to respond while the scale was on the screen, but that they could also respond after the scale disappeared.

In the scanner, the timing and instructions for each type of question appeared at the beginning of each block of 20 questions and a screen indicated the breaks between the runs. We blocked the items by questionnaire type to keep the procedure as close as possible to the actual taking of the STs outside the scanner and to allow any effects lasting for seconds after an item to carry over to items of the same questionnaire type most of the time. For the PCL and CES, each run had the full mST; for Neuroticism and Attachment, the 20 items in each run were a reasonable approximation of the number of items in the various tests used to measure these concepts.

We determined the timing of the presentations of the items using pilot testing to ensure participants had enough time to answer each question. We did not allow responses in the first 5 s to encourage participants to think about each question rather than simply guess and wait for the next trial. The jittered 3 s break after each question allowed a relaxed but not boring pace.

fMRI Acquisition and Analysis

Imaging, preprocessing, and basic analysis

Imaging was conducted on a 3 T GE MR750 MRI scanner (GE Healthcare, Waukesha, WI) with an eight-channel head coil. Head motion was minimized with foam pads, and participants wore earplugs to reduce scanner noise. The imaging sequence included a 3D plane localizer, followed by a T1-weighted structural acquisition and 3 runs of T2*-weighted (functional) acquisition. There was also a resting state scan at the beginning and the end of the run lasting 360 s. The resting-state runs were analyzed separately and will not be discussed further. Slice orientation was near-axial, parallel to the anterior–posterior commissure (AC–PC) plane. The T1-weighted anatomical images were 96 contiguous slices acquired with a high-3D fast inverse-recovery-prepared spoiled gradient recalled (SPGR) sequence, with repetition time (TR) = 3.22 ms, echo time (TE) = 8.2 ms, inversion recovery time (TI) = 450 ms, field of view (FOV) = 240 × 240 mm^2, 1.9 mm slice thickness, flip angle = 12°, voxel size = 0.9375 × 0.9375 × 1.9 mm, 256 × 256 matrix, and a parallel imaging with a selection factor of 2. The T2*-weighted echo-planar, functional images were sensitive to the blood oxygen level dependent (BOLD) signal. These were 34 contiguous slices acquired using a spiral-in sequence using sense imaging with a SENSE factor of 2, with TR = 2000 ms, TE = 30 ms, FOV = 240 × 240 mm^2, 3.8 mm slice thickness, flip angle = 70°, and voxel size = 3.75 × 3.75 × 3.8 mm.

Preprocessing and analyses of functional imaging data were conducted with Statistical Parametric Mapping software (SPM12; Wellcome Department of Cognitive Neurology, London, UK), along with locally developed MATLAB (Mathworks, Natick, MA) scripts. The first 3 volumes of each run were discarded to focus analyses on volumes acquired during steady-state equilibrium. Images were corrected for time difference of different slices in a volume, spatially realigned to the first volume to correct for motion, spatially normalized to the Montreal Neurological Institute (MNI) template using a 12-parameter affine model, and then spatially smoothed with an 8 mm full-width at half-maximum Gaussian kernel. A high-pass filter was included in every model to correct for scanner drift. Participants with between 3 and 5 mm of movement in 5% or fewer of their total scans were corrected with ArtRepair (http://cibsr.stanford.edu/tools/ArtRepair/ArtRepair.htm).

A general linear model approach was used to analyze the preprocessed data. In the first-level analysis, each questionnaire was modeled using a 6 s boxcar function convolved with a canonical hemodynamic response function. The boxcar function begins at the onset of each trial. No global normalization was used. Serial correlations between volumes due to noise and unmodelled neural activity were corrected using an autoregressive AR (1) model implemented in SPM12.

Separability of neural activity of the five tests

The separability of neural activity among the five questionnaires was established using Partial Least Squares Discriminant analysis (PLS-DA), using the libPLS package to classify whole-brain contrasts of each test versus baseline.

Classification models were assessed by performing leave-one-subject-out cross-validation (training on data from all but one subject and testing on the five contrast images for the left-out subject, repeating the procedure until all subjects are used for testing). This approach provides generalized estimates of performance and prevents overfitting. Because the classification problem involved five classes, a one-vs-all approach was adopted, wherein a classification model was developed for each of the five questionnaires and labels were assigned by identifying the model with the highest score. The number of latent variables in the PLS models was determined by minimizing the fivefold cross-validation error within the training folds (i.e., data for all but one subject). A model was then fit using all available training data and evaluated on the five questionnaires from the test subject. Inference on classifier performance was assessed across all contrasts using the observed number of correctly predicted contrasts and the binomial distribution B(115, .2), whereas inference on accuracy for individual questionnaires was performed using a B(23, .2) distribution. This use of binomial tests on the average leave-one-out classification accuracy is relatively unbiased because classifiers are stable across folds [Kohavi, 1995].

To evaluate the relationship between the semantic information contained in the questionnaires and the performance of the brain-based classifier, we compared the distribution of classification errors (across all unique combination of misclassification errors) to the semantic similarity of tests. Semantic textual similarity was computed between all pairs of test items using context-sensitive methods [Han et al., 2013]. Next, the average similarity of items from the 10 unique test pairs was computed and correlated (Spearman's rho) with the distribution of classification errors across subjects. Confidence intervals were estimated using bias accelerated bootstrapping. In this test, positive correlation coefficients indicate that tests with higher semantic similarity are more easily confused by the pattern classifier.

To map where increased brain activity led to classifications of each test type, bootstrap analysis was performed. In this procedure, the pool of contrasts was repeatedly resampled with replacement 5,000 times, with a new set of PLS models estimated on each iteration (with the number of latent variables fixed at 5). This procedure yielded a distribution of PLS regression coefficients at each voxel, which was converted to a z-score using normal approximation. Pearson correlation coefficients were computed to estimate the similarity between these classifier weight maps and univariate contrasts of each test versus baseline.

Comparisons of Attachment, Neuroticism, the PCL, and the CES to General Knowledge

Comparisons between each of the four questionnaires (Attachment, Neuroticism, PCL, and CES) versus General Knowledge were assessed separately in a second-level random-effect repeated-measure ANOVA. As visual and motor processes should be matched across conditions, this contrast should reveal differences in neural activity related to cognitive processes involved in taking these STs. The analysis was confined to voxels showing positive activation ($P < 0.05$, uncorrected) in the corresponding questionnaire compared to baseline to ensure that any differences reported were not driven only because of subtracting negative activation in the General Knowledge questionnaire. We calculated statistical thresholds by estimating the false positive rate using Monte Carlo simulation [Forman et al., 1995]. A simulation of 1000 iterations for each regression model produced a common threshold of cluster size = 41 with uncorrected threshold of $P < 0.001$ at each voxel to fulfill a corrected false positive rate of $\alpha < 0.05$, using AlphaSim implemented in the REST toolbox [Song et al., 2011, Updated September 2015]. Smoothness was estimated using the residuals from the ANOVA model.

To facilitate reverse inference of neuroimaging results, automated decoding was performed using the decoding tool on the Neurosynth website (http://neurosynth.org) [Yarkoni et al., 2011]. In this analysis, we computed the correlation between thresholded statistical maps for each of the four mSTs compared against General Knowledge and a series of automated meta-analytic maps for terms that frequently appear in the neuroimaging literature.

We used MNI coordinates throughout. Coordinates extracted from the literature in the Talairach space were transformed into the MNI space [Lancaster et al., 2007].

TABLE I. Correlations among the five modified standardized tests

	General Knowledge	Attachment	Neuroticism	CES
Attachment	0.26			
Neuroticism	0.25	0.76[c]		
CES	0.04	0.41[a]	0.51[b]	
PCL	−0.01	0.48[b]	0.54[b]	0.67[c]

[a] $P = 0.05$.
[b] $P < 0.05$.
[c] $P < 0.001$.

RESULTS AND DISCUSSION

Behavioral Results

The means and standard deviations for the individual subjects rating scales, which had response options of 1 to 4, were: General Knowledge 2.8, 0.3; Attachment 2.3, 0.5; Neuroticism 2.4, 0.4; the CES 2.5, 0.6; and PCL 1.6, 0.5. The reaction time differences in responding to the scales were not significant, $F(4, 110) = 1.55$, $P = 0.19$. The correlations among the means are shown in Table I.

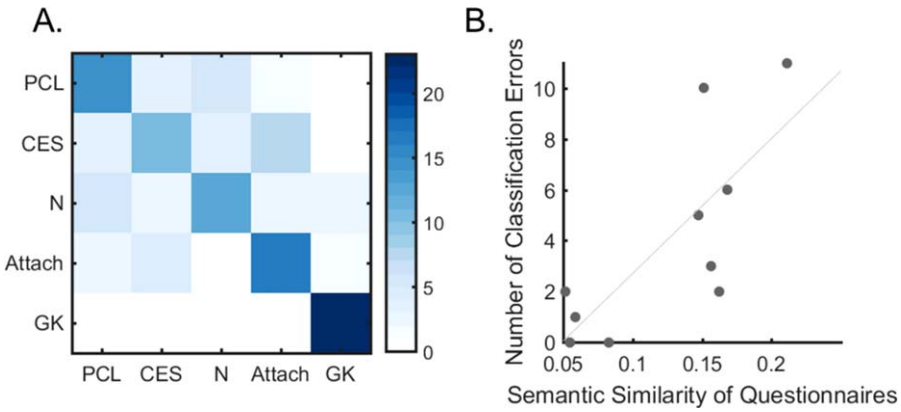

Figure 1.

Classifier performance. (A) Confusion matrix depicting the number of participants out of our total of 23 for which each test on the vertical axis was classified on the basis of neural activity in the other 22 participants as being the questionnaire on the horizontal axis. The off-diagonal row-wise elements indicate misses, and off-diagonal columns indicate false positives. PCL = Posttraumatic Stress Disorder Check List; CES = The Centrality of Event Scale; N = Neuroticism; Attach = Attachment; GK = General Knowledge. (B) Relationship between the semantic similarity of questions on modified standardized tests and the number of errors made in pattern classification. Each point reflects a unique pair of tests. The solid line depicts the best fitting least-squares estimate. [Color figure can be viewed at wileyonlinelibrary.com]

Separability of Neural Activity Among the Questionnaires

As shown in Figure 1a, the classifier performed well overall. Based on the neural activity of 22 of our 23 participants, we could classify correctly which of the five questionnaires the remaining participant was answering with an accuracy of 65%, where 20% is chance ($P < 0.0001$, binomial test). General Knowledge was the easiest to classify (100%, 95% CI [85.18% 100%], $P < 0.0001$), followed by Attachment (69.57%, [47.08% 86.79%], $P < 0.0001$), the PCL (60.87%, [38.54% 80.29%], $P < 0.0001$), Neuroticism (52.17%, [30.59% 73.18%], $P = 0.0001$), and the CES (43.48%, [23.19% 65.51%], $P = 0.0025$). There were no misclassification errors for General Knowledge and those for each of the four mSTs were fairly uniform among the other mSTs with a tendency for the PCL and Neuroticism and for the CES and Attachment to be more often misclassified with each other.

The correlation between the distribution of classification errors and the semantic similarity of test items indicated a high degree of similarity: *rho* = 0.7195, 95% CI [0.29 0.96] (Fig. 1b). These results confirm that the predictive information contained in patterns of neural activity was largely associated with the semantics of the tests, as opposed to unrelated factors that could be opportunistically leveraged in pattern classification. Given the separability of the brain activation and the correlation between the classification errors that did exist and the semantic similarity of the items, we turn to investigating how the four mSTs differ from the General Knowledge comparison task.

Activity of the Four mSTs Compared to General Knowledge

Differential levels of brain activation compared to General Knowledge were observed in three of the four mSTs. Details of these univariate comparisons are shown in Table II and Figure 2. The table and corresponding figure organize brain regions by their mSTs. These results are similar to those from pattern classification, as shown in Figure 3.

To provide a post-hoc assessment of brain networks engaged during the mSTs, we quantified the degree of overlap between the four comparisons against General Knowledge and previously characterized functional networks (Table IV). The proportion of activated voxels was computed using binary masks from two meta-analyses of Theory of Mind [Carter and Huettel, 2013; Mar, 2011] and two delineations of the default network using resting-state fMRI [Shirer et al., 2012; Yeo et al., 2011]. In the two default network masks, Attachment had 38 and 50% of voxels activated; in the two mentalizing network masks, 64 and 91%. The CES and PCL voxel activations were similar to each other, with roughly half the voxels activated as those reported for Attachment. Neuroticism had <1% of voxels activated in any of the four networks. Thus, brain regions engaged by the Attachment mST reliably activate the majority of areas that are typically observed during Theory of Mind tasks (and to a lesser degree the default network).

To provide further post-hoc characterization of our results, we conducted a data-driven assessment of terms associated with these maps from Neurosynth. The results are provided in Table III. Contrast maps and the full assessment of terms can be explored in full at the NeuroVault website (neurovault.org/collections/DLPYYJTY).

TABLE II. Comparisons of the modified standardized tests to General Knowledge

Standardized Test	Brain region	Hemisphere	MNI coordinates			Z score	Voxels
			x	*y*	*z*		
Attachment > General Knowledge	Dorsal medial PFC	Bilateral	−4	53	34	5.76	379
	Dorsal medial PFC	Bilateral	0	56	23	5.41	
	Dorsal medial PFC	Bilateral	−8	56	11	5.39	
	Posterior cingulate/precuneus	Left	−8	−49	38	6.49	49
	Angular gyrus	Left	−53	−60	30	5.65	203
	Middle temporal gyrus	Left	−49	−56	19	5.12	
	Middle temporal gyrus	Left	−41	−38	0	3.61	
	Middle temporal pole	Left	−49	8	−34	4.79	48
	Middle temporal gyrus	Left	−53	4	−27	4.66	
	Inferior temporal gyrus	Left	−45	0	−38	4.53	
CES > General Knowledge	Paracentral lobule	Bilateral	−15	−26	72	4.99	265
	Paracentral lobule	Bilateral	−4	−23	72	4.64	
	Supplementary motor area	Bilateral	8	−15	72	4.62	
	Angular gyrus	Left	−53	−60	30	4.11	48
	Angular gyrus	Left	−45	−56	42	3.42	
PCL > General Knowledge	Paracentral lobule	Bilateral	−4	−19	72	4.83	345
	Mid cingulate cortex	Bilateral	0	−19	38	6.48	
	Mid and posterior cingulate	Bilateral	−8	−41	34	5.13	
	Angular gyrus	Left	−45	−56	42	4.86	73
General Knowledge > Attachment	Insula	Right	34	23	0	5.66	43
	Inferior frontal gyrus, opercular	Left	−41	8	27	4.94	92
	Inferior frontal gyrus, triangular	Left	−45	30	23	4.50	
	Middle occipital cortex	Left	−38	−83	30	4.25	46
	Superior parietal cortex	Left	−26	−68	46	4.16	
General Knowledge > CES	Inferior frontal gyrus, triangular	Left	−49	30	19	5.99	325
	Precentral gyrus	Left	−45	8	30	5.70	
	Inferior frontal gyrus, triangular	Left	−49	19	30	5.15	
	Hippocampus	Bilateral	−4	−15	−19	5.72	106
	Thalamus	Bilateral	−4	−23	−8	5.18	
	Parahippocampal gyrus	Bilateral	11	−23	−15	4.71	
	Insula	Right	34	23	0	5.52	69
	Middle frontal gyrus	Right	49	30	23	5.27	165
	Inferior frontal gyrus, opercular	Right	41	15	30	4.77	
	Inferior temporal gyrus	Left	−45	−56	−8	4.89	150
	Fusiform gyrus	Left	−38	−45	−23	4.67	
	Fusiform gyrus	Left	−41	−56	−23	4.55	
	Supplementary motor area	Bilateral	−4	19	49	4.64	85
	Supplementary motor area	Bilateral	4	15	57	4.03	
General Knowledge > PCL	Insula	Left	−30	26	0	5.68	64
	Insula	Right	30	23	−4	5.56	65
	Thamalus	Bilateral	0	−26	−8	5.11	88
	Prarahippocampal gyrus	Bilateral	−4	−11	−19	4.89	
	Parahippocampal gyrus	Bilateral	11	−26	−15	4.27	
	Middle frontal gyrus	Right	49	30	23	4.24	61
	Inferior frontal gyrus, triangular	Right	41	23	27	3.85	
	Precentral gyrus	Left	−41	8	30	4.23	84
	Inferior frontal gyrus, triangular	Left	−49	30	19	3.69	
	Inferior frontal gyrus, opercular	Left	−53	23	34	3.58	
General Knowledge > Neuroticism	Fusiform gyrus	Left	−26	−34	−23	5.81	67
	Fusiform gyrus	Left	−38	−41	−30	3.98	
	Insula	Right	30	23	−4	5.53	51
	Inferior frontal gyrus, orbital	Right	26	34	−8	3.69	
	Middle frontal gyrus, orbital	Right	34	38	−15	3.54	
	Middle occipital cortex	Right	38	−71	38	5.53	111
	Insula	Left	−30	23	0	5.15	54

TABLE II. (continued).

Standardized Test	Brain region	Hemisphere	MNI coordinates			Z score	Voxels
			x	*y*	*z*		
	Inferior frontal gyrus, orbital	Left	−30	34	−11	4.75	
	Inferior frontal gyrus, orbital	Left	−26	26	−15	3.37	
	Middle occipital cortex	Left	−38	−79	30	5.14	151
	Superior parietal cortex	Left	−26	−75	53	4.85	
	Inferior temporal gyrus	Left	−56	−56	−8	5.03	86
	Middle frontal gyrus	Right	49	30	23	4.66	133
	Inferior frontal gyrus, opercular	Right	41	19	30	4.30	
	Inferior frontal gyrus, opercular	Right	53	11	34	4.08	
	Inferior frontal gyrus, triangular	Left	−45	30	19	4.65	170
	Precentral gyrus	Left	−41	11	30	4.53	
	Parahippocampal gyrus	Bilateral	−4	−11	−19	4.57	65
	Hippocampus	Bilateral	4	−11	−15	4.23	
	Thalamus	Bilateral	0	−26	−8	4.01	
	Dorsal medial PFC	Bilateral	4	30	46	4.46	57
	Supplementary motor area	Bilateral	−8	23	46	4.12	

Several patterns are consistent with the results of activation in the mentalizing and default networks, but from a very different form of analysis. The terms "theory of mind," "mentalizing," "mind," and "mental states" appear only for Attachment. The terms "default mode," "default," "mode," "mode network" appear for Attachment, the CES, and the PCL. Neuroticism does not contain any of these terms but does include terms that may be consistent with anxiety in that they focus on not acting including "inhibit," "stop signal," "cognitive control," and "inhibitory." The anatomical terms are generally consistent with the activity noted in Figure 2. In general, the terms associated with these contrast maps, though not specified *a priori*, are reasonable and do not point to artifact-based explanations of our basic findings.

The next section is organized by individual brain regions, noting similarities and differences in each brain region among the mSTs and interpreting activity in each brain region using meta-analyses of neural activity related to particular areas or tasks (see Table II for a list of clusters). We start with the angular gyrus (AG) because its interpretation led us to appreciate the importance of specific Theory-of-Mind (ToM) processes that helped us to interpret other regions.

Individual Areas Active for the Four mSTs Compared to General Knowledge

Angular gyrus

The AG appears in Attachment, the CES, and the PCL (see the left lateral column of Fig. 2 for these regions and to place the area shown in Fig. 4 in context). Figure 4A shows the overlap of Attachment, the CES, and the PCL within the AG and surrounding area in more detail. We chose to probe activations with findings from meta-analyses of neural activation during ToM, or mentalizing tasks, because they often activate the AG and because such tasks involve the ability to understand one's own and other people's mental states and actions. Thus, such tasks have direct relevance to our mSTs. Figure 4B includes maps of activity for ToM from two meta-analyses—Carter and Huettel [2013] and Mar [2011]—that fell within the AG. As Figure 4B shows these ToM areas fell within the union of activity in our three individual differences measures. Carter and Huettel [2013] also produced reverse-inference maps for "intention," "mentalizing," and the more conceptually diffuse term of "social," not shown, that fell within the ToM area in the AG. Carter and Huettel [2013] and Mar [2011] measured other aspects of social cognition that did not overlap with the union of activity shown in Figure 4B, providing an indication of the specificity of the processes [Schaafsma et al., 2015]. For Carter and Huettel, these included eye gaze, facial expression and biological motion; for Mar, this included narrative comprehension, which did overlap with Attachment, but mainly in the temporal lobe rather than in the AG.

To explore more broadly the tasks that elicit activity in the AG to specify the tasks that commonly activate and do not activate the regions we found, we used Seghier's [2013] review of meta-analyses of the localization of multiple functions in the AG. It includes all meta-analyses he could find in this area independent of the comparisons they investigated. Based on 16 meta-analyses, he found 23 categories of tasks that had consistent activation within the AG. Two peaks, story-based and non-story-based ToM, were from Mar [2011] and are shown in the Figure 4B by their extent. Of the remaining 21 areas, only two, Spreng et al. [2009]'s ToM and Sperduti et al.'s [2011] external intention, showed peaks of activity within the boundary of the union of activity in our three tasks. External intention, which has a peak dorsal to the peaks from the other meta-analyses shown in Figure 4B, refers to tasks in which participants' finger or hand movements indicated the control of movement involved in an external agent; this seems

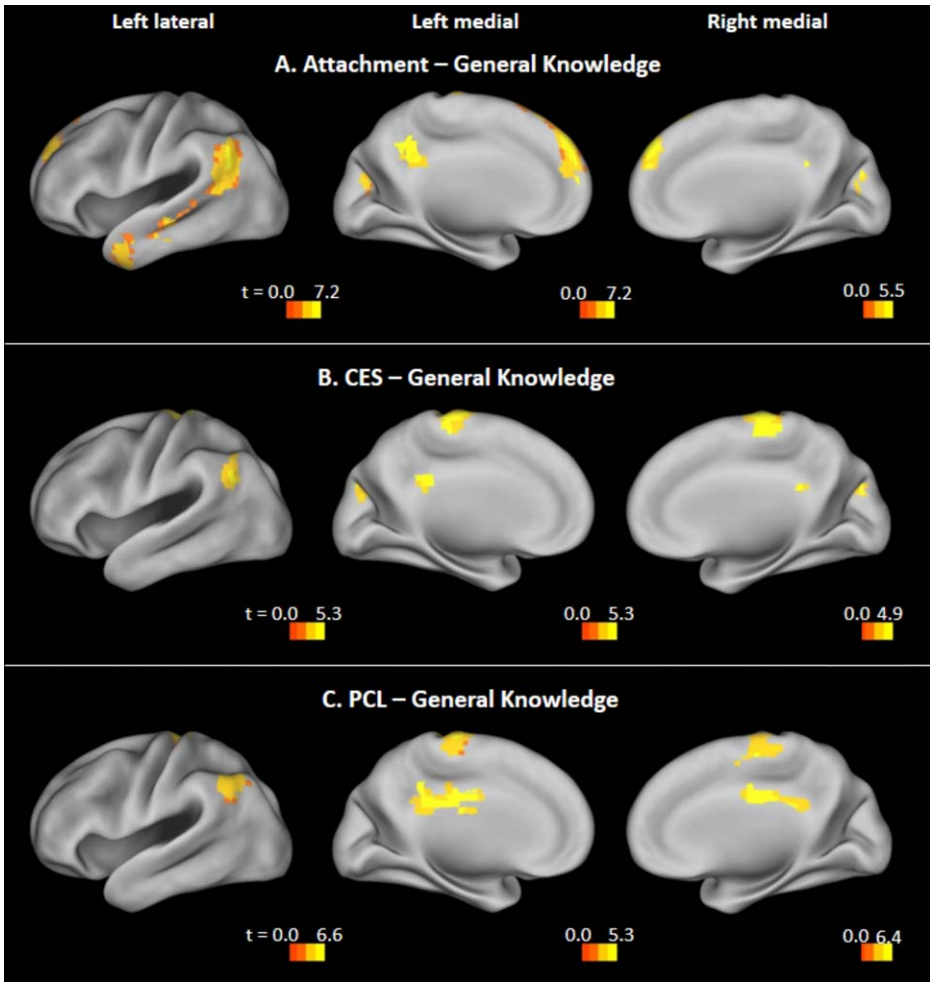

Figure 2.

Whole-brain activations of (A) Attachment, (B) CES, and (C) PCL, versus General Knowledge ($N = 23$). Activations were thresholded at $P < 0.001$ at the voxel level. Only clusters with more than 41 voxels are shown in the figure.

unrelated to our tasks. The remaining 19 peaks that did not lie within the boundary contained a variety of tasks involving the default network, laboratory episodic memory retrieval and recollection, autobiographical memory, semantic memory, and visual-spatial tasks. Similarly, a review of individual studies involved in social cognition [Van Overwalle, 2009] found areas involved in action and outcome monitoring, human faces, bodies, movement, and the mirror system did not lie within the boundary shown in Figure 4B, whereas those involved in ToM did. Thus, the ToM tasks that overlapped or had peaks within the union of activity in our three tasks are focused conceptually. However, it is possible that activations indicated by peaks and areas indicated by their extent near the boundary of our AG activation extended into or were active but just under our statistical threshold.

Because most of our activations are in areas previously found in ToM studies, we did a search for meta-analyses of ToM using fMRI, which added one paper, Bzdok et al. [2012]. This meta-analysis included three related categories: ToM tasks that required the prediction of the thoughts, intentions, and future actions of other people, morality tasks that required participants to make appropriateness judgments of the actions of one individual towards others, and empathy tasks that were intended to elicit the conscious experience of someone else's affective state. As shown in Figure 4B, all three had peaks of activity within the union of activity in our three tests. Also shown are two peaks from Denny et al. [2012] with similar coordinates for judgments about others either greater than a baseline or greater than judgments about the self.

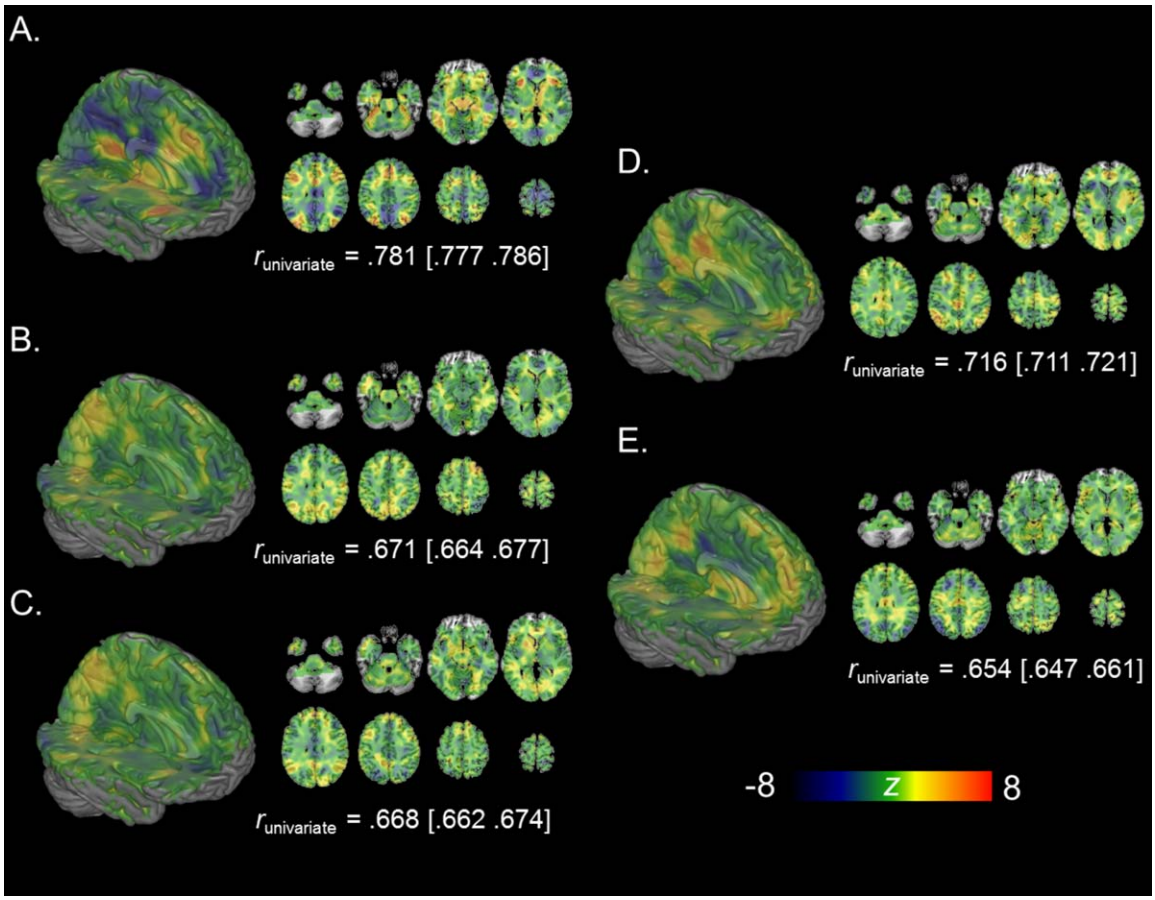

Figure 3.

Whole-brain multivariate patterns of activation predictive of each test. (A) General Knowledge; (B) Centrality of Event Scale; (C) Attachment; (D) PTSD Checklist; (E) Neuroticism. Parametric maps reflect *z*-scores computed from bootstrap resampling of classification coefficients at each voxel. Text overlays show the Pearson correlation coefficient and 95% confidence interval between each classification weight map and the corresponding univariate contrast of each test against the average of all others.

Thus, the area in the AG that was active in three of our mSTs appears often in tasks related to ToM. These ToM tasks include many neural processes and activate many areas outside of the AG, but there is a clear and fairly localized overlap between the area in the AG active during our mSTs tests and the activity in the AG caused by tasks related to ToM.

Outside the AG, areas active in ToM meta-analyses include most activity we note in the posterior cingulate cortex (PCC), dorsomedial prefrontal cortex (dmPFC), and the anterior temporal lobe (TL). This unexpected finding, which is described in more detail in the sections that follow, suggests possible ToM and social processes in the STs we used. It should also be noted that the meta-analyses we report are based on different search, selection, and analysis procedures applied to the same expanding corpus of research. Nonetheless, this corpus is substantial enough to produce reliable results with the more recent meta-analyses each analyzing over 60 studies.

Cingulate cortex

The cingulate cortex is active in Attachment and the PCL (Fig. 2). Adopting the anterior, mid, and posterior division of the cingulate [Apps et al., 2013; Palomero-Gallagher et al., 2009; Vogt et al., 1995], all cingulate activations occur in the mid and posterior cingulate on the gyrus, which have different Brodmann's areas and thus different cytoarchitectonics than the anterior and the sulcal regions.

The PCC is part of the mentalizing network outlined by Mar [2011], which includes cingulate cortex that overlaps with our Attachment activity for both story-based and non-story-based ToM. Moreover, peak coordinates from

TABLE III. Top 10 terms from automated decoding of contrasts of the modified standardized tests against General Knowledge

	Attachment		CES		N		PCL	
Rank	Term	*r*	Term	*r*	Term	*r*	Term	*r*
1	Medial prefrontal	0.355	Precuneus	0.309	Inhibit	0.176	Precuneus	0.271
2	Theory mind	0.321	Foot	0.247	Cingulate	0.117	Posterior cingulate	0.260
3	Mentalizing	0.302	Posterior cingulate	0.216	Stop signal	0.113	Foot	0.229
4	Mind	0.299	Cortex precuneus	0.188	Middle cingulate	0.112	Cingulate	0.190
5	Precuneus	0.298	Precuneus posterior	0.154	Anterior posterior	0.112	Default mode	0.177
6	Mental states	0.287	Default mode	0.137	Cognitive control	0.107	Mode	0.170
7	Posterior cingulate	0.286	Arm	0.137	Group	0.102	Cortex precuneus	0.169
8	Default	0.274	Limb	0.136	Cingulate cortex	0.096	Default	0.168
9	Medial	0.265	Default	0.132	Posterior cingulate	0.088	Precuneus posterior	0.154
10	Temporo parietal	0.264	Mode	0.131	Inhibitory	0.085	Mode network	0.154

Note: *r* = Pearson's correlation coefficient; CES = centrality of events scale; N = Neuroticism; PCL = PTSD checklist.

Bzdok et al.'s [2012] meta-analysis for Morality ($MNI_{x,y,z}$: 0, −56, 34) and Spreng et al.'s [2009] meta-analysis for ToM (−4, −52, 31) are both located in this overlapping area. Because the PCC is also a key node in the default network, it has generally been assumed to be involved in thought that occurs when the mind is free to wander. These thoughts include autobiographical memories and future events, which are related to the type of cognitive processes evoked by our mSTs [Buckner et al., 2008; Buckner and Carroll, 2007; Raichle, 2015]. It has also been suggested that the PCC is central to cognitive processes involved in being attached to or caught up in one's memory [Brewer et al., 2013]. This can be seen as an alternative way of describing the CES in that the CES is a scale measuring the perceived centrality of the memory of a negative event to one's life, which correlates about 0.55 with PTSD symptom severity [Rubin et al., 2014]. The mid cingulate was active in the PCL. The gyrus of the mid cingulate cortex is involved in a variety of functions, but especially those related to social cognition [Apps et al., 2013].

Paracentral lobule

The paracentral lobule is active bilaterally in the CES and PCL. We could not find any meta-analyses for this region. Several studies of PTSD using a variety of tasks found areas that overlap our paracentral gyrus activity. Shin et al. [2001] found an area in the left paracentral lobule (peak: −10, −31, 43) that was more active in Vietnam combat participants with PTSD than those without. Garrett et al. [2012] found an area in the left paracentral lobule (peak: −13, −27, 46, 52 voxels) in youth with PTSD symptoms and a history of interpersonal trauma compared to controls. Hennig-Fast et al. [2009] found an area in the paracentral lobule (peak: −2, −35, 50, 348 voxels) that was more active for traumatized compared to nontraumatized participants. This leaves us with three interpretations. The first is that we are observing effects related to motor activity. This is unlikely because in all five tests the motor activity is matched and reaction times did not differ among the five tests ($F(4, 110) = 1.5$, $P = 0.2$). The second interpretation is that this area is affected by PTSD or the effect of specific traumas in a way and under conditions we do not yet understand. However, as described in the Relation of Mean Scores and Activity section, there was no effect of the PCL score or CAPS on our results. The third interpretation is that the paracentral lobule activity is a chance finding or one with a small effect size that occasionally is observed.

Dorsomedial prefrontal cortex

The dmPFC was active in Attachment. Figure 2A shows the dmPFC activation in relation to all other activity in Attachment. Figure 5 shows the overlap of the Mar [2011]

TABLE IV. Percentage of activated voxels in cortical networks for contrasts of the modified standardized tests against General Knowledge

Cortical network	Publication	Attachment	CES	*N*	PCL
Default network	Yeo et al. [2011]	0.385	0.150	0.006	0.189
	Shirer et al. [2012]	0.498	0.248	0.008	0.296
Mentalizing network	Carter and Huettel [2013]	0.635	0.107	0.000	0.108
	Mar [2011]	0.909	0.425	0.000	0.404

Note: CES = Centrality of Event Scale; N = Neuroticism; PCL = PTSD checklist.

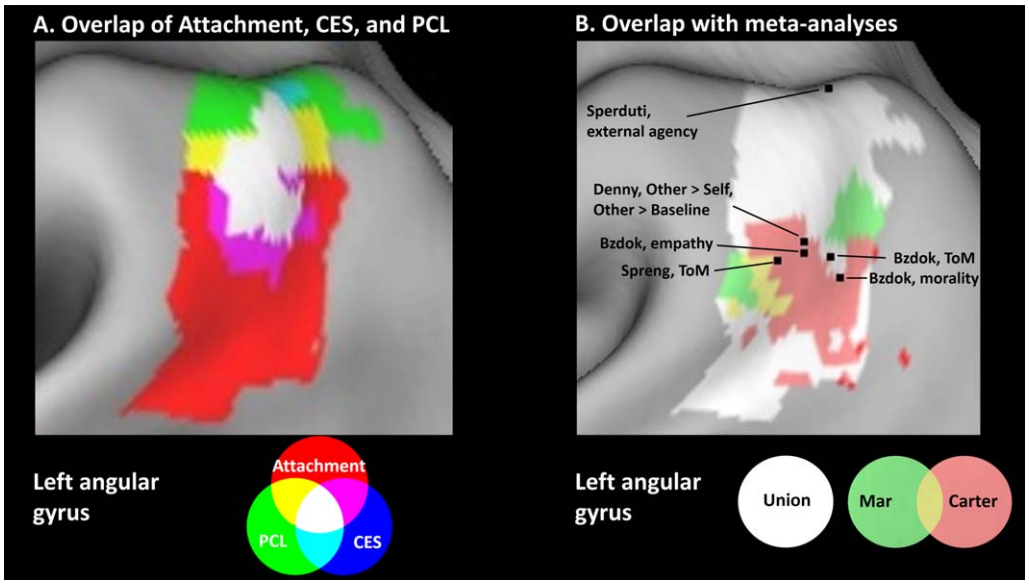

Figure 4.

Activation of the left AG compared to General Knowledge. (A) Left AG activations from Attachment, CES, and PCL. (B) Overlap of the left AG regions shown in two meta-analyses [Carter and Huettel, 2013; Mar, 2011] and the left AG region activated by the union of Attachment, the CES, and the PCL in this study. Peak activities are shown for other meta-analyses.

and Carter and Huettel [2013] meta-analyses. It also shows peaks from other meta-analyses that had peaks within the Attachment activation. Denny et al. [2012], unlike other meta-analyses, distinguish between judgments about the self versus others, which is especially relevant here because Attachment stresses how the self relates to others. When self and other judgments are compared as shown in Figure 6A,B, other > self overlaps with Attachment, whereas self > other does not but is uniformly ventral to Attachment. For regions in which both self and other judgments are greater than baseline, as shown in Figure 6C, Attachment activity is mostly within that of the meta-analysis regions, except for a dorsal PFC region on the right. Thus, Attachment likely relates both to self and other judgments but is best described by regions in which judgments of others produce more activity than judgments of self. Figures 4 and 5 offer support to the activity from the Attachment mST being related to ToM judgments, especially those about others.

Left anterior middle and inferior temporal gyrus

There are five peaks of activity in two clusters in the left temporal gyrus, all for Attachment. Strong support for the involvement of the anterior temporal lobes in social cognition comes from lesion studies in animals, neuropsychology, and neuroimaging [Olson et al., 2013; Rice et al., 2015; Wong and Gallate, 2012]. This typically includes memory for other people and their relationships, social language, and social behavior. Most reviews find little evidence for laterality effects in the anterior temporal lobes in general or specifically for social cognition [Rice et al., 2015; Wong and Gallate, 2012] or for more specific localization within the anterior temporal lobes; however, some reviews support more precise localization. Specifically, these reviews found a region that overlaps with ours for a variety of ToM tasks that did not include the paradigm of understanding stories about ToM problems [Mar, 2011; Olson et al., 2007, 2013]. Thus, the strong left lateralization and lack of activity in the superior temporal gyrus that we observe are not supported by the reviews but seem specific to the current results. The left lateralization in the anterior temporal lobes may be driven by general left lateralized cognitive processes such as language, ToM, and autobiographical memory [Kim, 2012]. For the other left lateral activations, as reviewed, the left lateralization is consistent with the literature.

General Issues Regarding Activity of the Four mSTs Compared to General Knowledge

Relation of mean scores and activity

To investigate whether the mean scores on the four mSTs affected the activity that occurred when participants answered test items during the scans, we correlated the average of each participant's ratings over all 60 items rated for each of the mSTs with the activity in the clusters

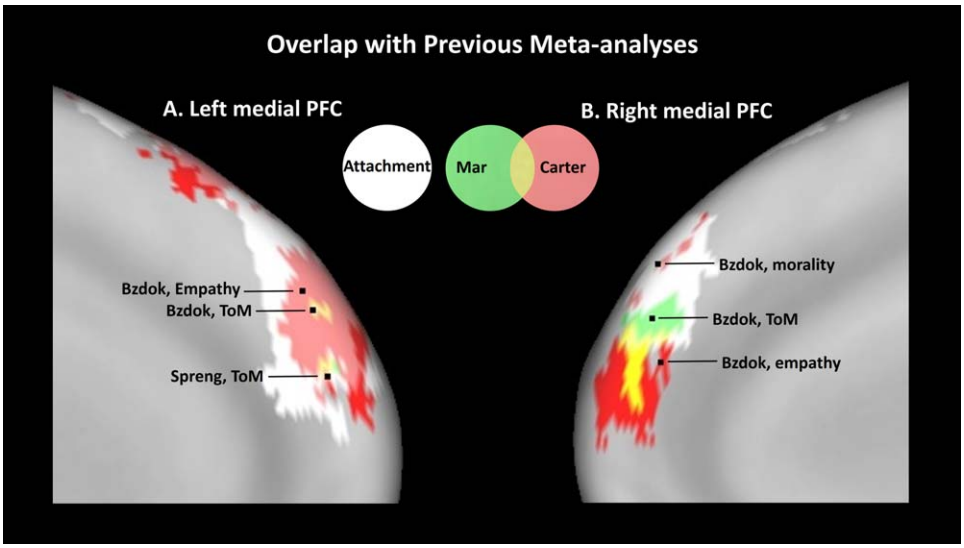

Figure 5.

Activation of the mPFC for Attachment compared to General Knowledge and regions from two meta-analyses [Carter & Huettel, 2013; Mar, 2011]. (A) Effects in left medial PFC. (B) Effects in right medial PFC. The more saturated areas, or red and yellow, occur where the Carter and Huettel meta-analysis and its overlap with the Mar meta-analysis do not overlap with Attachment. Peak activities are shown for other meta-analyses Bzdok et al. 2012; [Spreng et al., 2009]. PFC = prefrontal cortex.

shown for that mST. Of the 8 clusters that have a mST > General Knowledge in Table II, two were significant at the uncorrected 0.05 level. To see if PTSD symptom severity affected the activity shown, we correlated the participants' CAPS scores from prior to their scans with the activity of each of these 8 clusters. Again, two correlations were significant at the 0.05 level. None of these four clusters survived correction for multiple comparisons. If the mean ratings of any of the individual participants on any of the four mSTs or the CAPS had effects on processes involved in producing the clusters shown in Table II, the effects were too small to observe here. Differences in neural processes might be observed with more statistical power, but there is no a priori theoretical reason to assume that people who score at different levels on an ST use or do not use different neural processes; they could be using either the same or different processes to make different responses. In sum, to a first approximation, the neural correlates of the cognitive and affective processes involved in completing the four mSTs are independent of scores they produce and of PTSD symptom severity.

Alternatives to the use of a single comparison task

After demonstrating that the different questionnaires elicited differentiable brain activation through pattern classification, all the remaining analyses reported, which described and interpreted their activations, were based on comparisons to the single comparison task of General Knowledge. For these analyses, we restricted the analysis to voxels showing positive activation ($P < 0.05$, uncorrected) in the corresponding questionnaire compared to baseline. Nonetheless, given that this is the first study of its kind, we wanted to ensure the selection of this comparison task was not primarily responsible for our results. Therefore, in addition to the qualitative observation that our four mSTs, while often overlapping, activated different neural regions, we conducted two additional lines of analysis. First, we compared our mSTs to each other rather than comparing each mST to the same baseline. The purpose of these comparisons is to show that differences occur using a variety of baselines, not to try to interpret or speculate about each of the differences. Table V includes these comparisons grouped by the less active baseline cluster. As shown in Table V, when our mSTs were compared to each other instead of General Knowledge, many areas showed differential levels of activation; all mSTs had at least one area greater than and at least one area less than another mST. These differences indicate that it is unlikely that our results were caused by using General Knowledge as the baseline.

Second, we performed an analysis examining the direction of the signal change in regions that exhibited differences between mSTs and General Knowledge to determine whether deactivation in General Knowledge was driving these findings. The left AG and the PCC are significantly active when comparing General Knowledge with Attachment, the PCL, and the CES as shown in Figure 7A,B. However, the PCC is a key node in the default network [Buckner et al., 2008] and so it is expected that it will be deactivated in General Knowledge, a task involving semantic memory, because default mode network regions are typically deactivated during task performance, unless

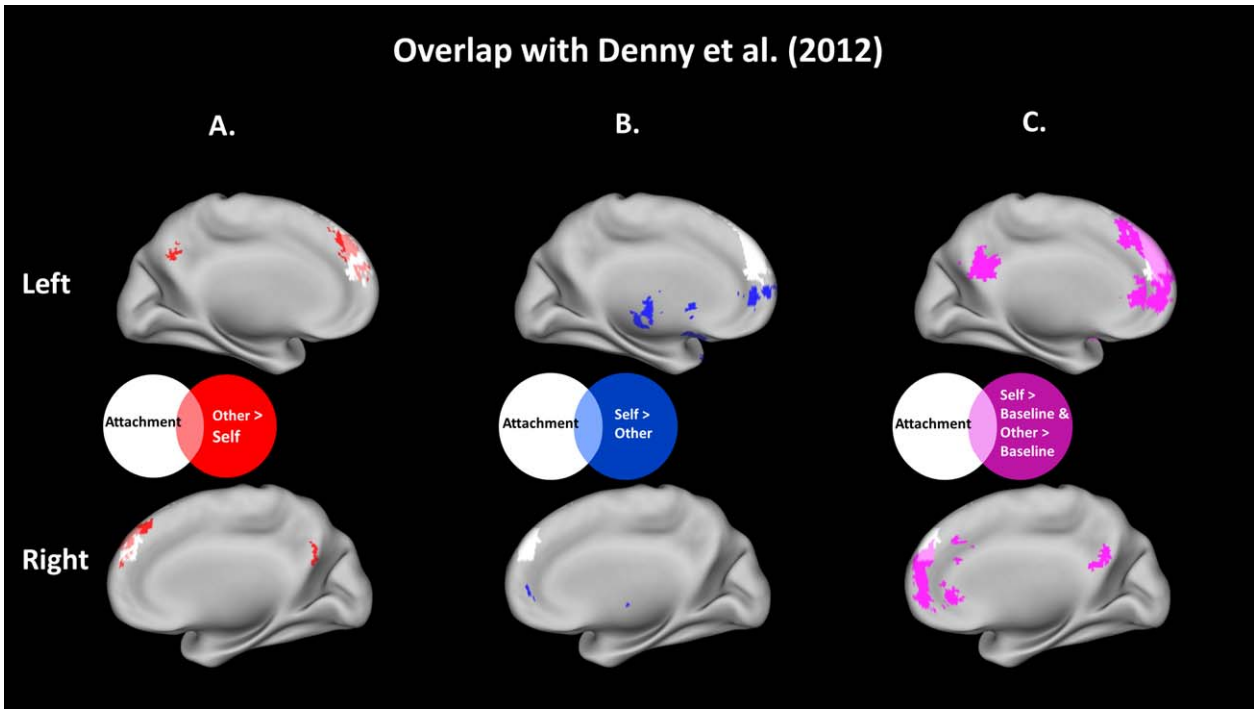

Figure 6.
Activation of the mPFC for Attachment and regions from meta-analysis by Denny et al. [2012]. (A) Effects of other versus self show overlap with Attachment in the mPFC. (B) Effects of self versus other do not show overlap with Attachment. (C) Conjunction of effects for self versus baseline and other versus baseline shows overlap with Attachment in the mPFC.

the task involves self-referential thought. This is what we observe. One could argue that this below baseline activity is an artifact that is driving our results or that, as would be expected from theoretical considerations, our standardized tests differ from tests of semantic memory. Given the other evidence presented, especially the activity in the PCC during comparisons of these three mSTs with Neuroticism, we prefer the latter interpretation for the PCC. For other key areas such as the dmPFC and left temporal pole, shown in Figure 7C, there is little General Knowledge deactivition to subtract.

The limited activity for Neuroticism

Three of the four mSTs showed considerable activity above the General Knowledge comparison: the PCL, CES, and Attachment tests. As shown in the previous section, when we compared Neuroticism to each of the other three mSTs using the same methods and statistical thresholds as we did for the General Knowledge baseline, there was activity only in the occipital cortex when compared to the CES. Thus, although Neuroticism is key to understanding personality, is the main personality dimension in many clinical disorders, and is associated with activity in brain regions related to anxiety that could reasonably be expected to be differentially active when items that measure it were being rated, we did not find effects for it. One reason could be that the neural changes are smaller in magnitude for Neuroticism and so require more power than was needed for the other mSTs.

GENERAL DISCUSSION

Summary

To our knowledge, this is the first attempt to study the neural activity underlying the processes used in taking STs. We showed differential brain activity across our four mSTs and our control task using a pattern classification analysis that clearly demonstrated differences that cannot be attributed simply to reading and making multiple-choice decisions. Our methods and analyses were rigorous, allowing us to exclude many alternative explanations of the activity we found. We had no studies of test-taking from which to draw predictions regarding anatomical localization, but activity could be related to existing meta-analyses of fMRI studies of processes related to areas identified for our mSTs. There were meaningful interpretations based on meta-analyses for all areas except for the paracentral lobule, for which we could not find a meta-analysis.

TABLE V. Comparisons of the modified standardized tests with alternative comparison tasks

	>Neuroticism	>Attachment	>CES	>PCL
<Neuroticism			Bilateral occipital cortex (left: 76 voxels, peak: −23, −101, −4; right: 49 voxels, peak: 38, −90, −4)	
<Attachment	Bilateral inferior parietal cortex (left: 189 voxels, peak: −45, −56, 27; right: 48 voxels, peak: 41, −71, 42) Posterior cingulate cortex (PCC, 108 voxels, peak: 0, −56, 38) Bilateral dlPFC, extending into dmPFC (left: 209 voxels, peak: −23, 23, 53; right: 135 voxels, 26, 23, 49).		dmPFC (73 voxels, peak: −4, 53, 30), striatum, thalamus, and brain stem (104 voxels, peak: −4, −23, −15) vlPFC (61 voxels, peak: −49, 15, 15)	Striatum, thalamus, and brain stem (117 voxels, peak: −15, 8, 15) dmPFC (62 voxels, peak: −8, 15, 68)
<CES	Bilateral inferior parietal cortex (left: 304 voxels, peak: −38, −75, 42 right: 144 voxels, peak: 41, −71, 38) PCC extending into precuneus (432 voxels, peak: 0, −34, 42) Bilateral dlPFC (left: 82 voxels, peak: −23, 23, 53; right: 80 voxels, peak: 34, 15, 57).			
<PCL	Bilateral inferior parietal cortex (left: 297 voxels, peak: −45, −56, 46; right: 170 voxels, peak: 41, −71, 42) PCC (155 voxels, peak: 0, −30, 42), PCC extending into precuneus (246 voxels, peak: 0, −53, 19) Bilateral dlPFC (left: 98 voxels, peak: −23, 23, 53; right: 45 voxels, peak: 26, 23, 46) Inferior temporal gyrus (47 voxels, peak: −56, −41, −15) Orbitofrontal cortex extending into anterior cingulate cortex (95 voxels, peak: −23, 60, 19)	PCC (58 voxels, peak: 0, −30, 42)		

An important finding was the degree to which areas found in meta-analyses of ToM coincided with the activity we found for our mSTs. Although the activity compared to General Knowledge in three of our mSTs was different, the PCL, CES, and Attachment had most of their activity in regions that overlapped with the mentalizing network, and only in a fairly focused subset of the mentalizing network. This relationship was also supported by decoding analyses performed with Neurosynth, which revealed a positive correlation between the Attachment contrast and meta-analytic maps of related terms including "theory (of) mind," "mentalizing," and "mind." These results stress the mentalizing and social nature of this test. Conversely, it provides a new way to view the default and mentalizing networks in terms of STs related to PTSD, such as the PCL, CES, and Attachment, tasks that had not been used to probe it [Schaafsma et al., 2015]. In particular, not only are these networks related to social and philosophical ToM concepts [Andrews-Hanna et al., 2014], but also to test-taking in terms of answering questions probing clinical and individual differences concepts as shown here. These alternative views of the same neural substrates broaden the behavioral scope of established cortical networks and offer a way to better understand them by using converging evidence. One important question for future work is whether the default and mentalization networks would show similar activation associated with answering other psychometric tests or whether this activation is specifically associated to the tests included here.

The STs we used were developed with great care, validated over many behavioral studies, and have large literatures relating them to other tasks and behavior. Interpretations of the processes people use in completing them are informed by those literatures. Both the overlap of activity in the mPFC for Attachment and the activity shown in the mPFC in meta-analyses in which there was more activity in tasks about others compared to the self [Denny et al., 2012] (Fig. 7) are examples. Moreover, the ToM tasks in the meta-analyses that coincided were ones that had obvious relevance to their mSTs, whereas less relevant ToM tasks did not coincide. We do not expect ToM tasks to be relevant to all STs, but for the included mSTs and for other clinical disorders this observation warrants further investigation.

The results reveal a considerable amount about processes involved in answering STs related to PTSD. Areas active while taking the PCL included the AG and PCC.

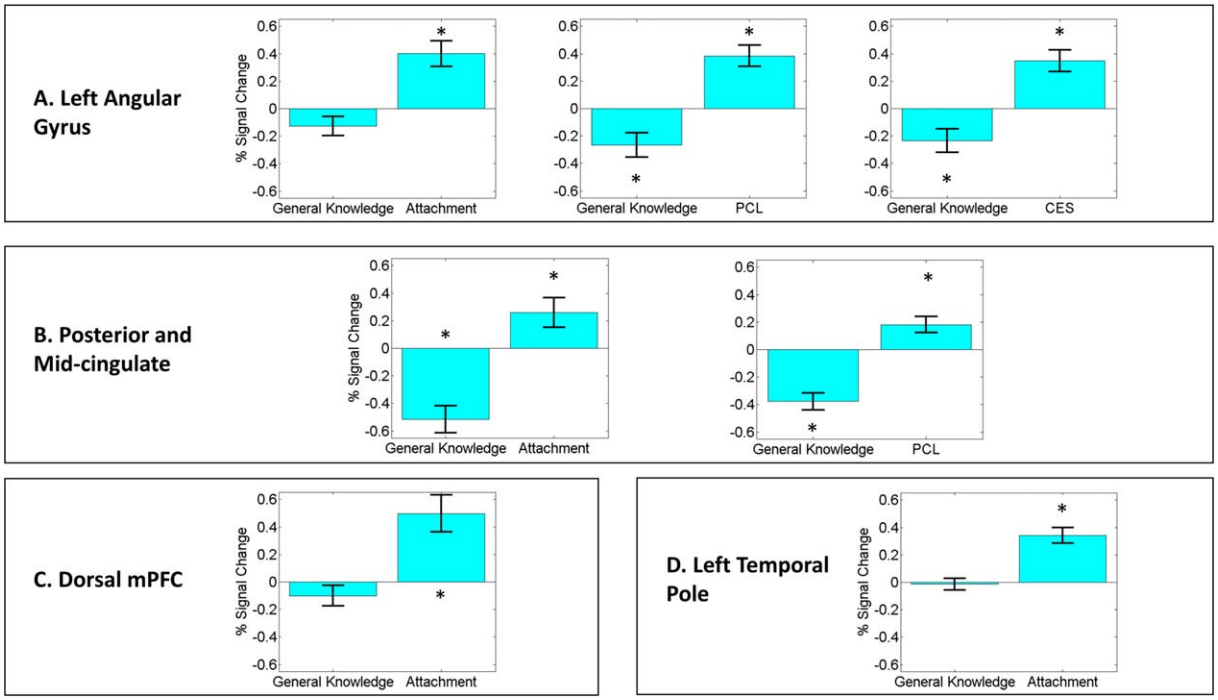

Figure 7.

Percent signal change in regions of interest. Each bar plot depicts percent signal change of the brain region showing significant activation when comparing the corresponding questionnaire with General Knowledge. (A) Signal change in left angular gyrus. (B) Signal change in posterior and mid-cingulate cortex. (C) Signal change in the dorsomedial prefrontal cortex. (D) Signal change in the left temporal pole. The regions of interest were functionally defined in the corresponding contrasts against General Knowledge. Error bars represent SEM. [Color figure can be viewed at wileyonlinelibrary.com]

These areas overlap partially with activity during answering the CES and Attachment, but the PCL had additional mid cingulate activity. In contrast, the CES and Attachment had activity in areas that the PCL did not. Such partial overlap holds promise for finding areas that are common to and unique to these STs, suggesting hypotheses about common and unique processes. It also indicates a method and analysis that may be useful in domains other than PTSD. For now, we note that our results suggest the importance of social processes in answering individual differences tests related to PTSD; independent evidence exists for social processes in PTSD itself [Charuvastra and Cloitre, 2008; Ogle et al., 2015].

LIMITATIONS

While introducing a new approach with potentials for significantly expanding our knowledge of psychometric test-taking, this study also involves a number of limitations for future research to consider. First, the four mSTs consisted of questions that were aggregated from existing scales and adjusted to fit the fMRI environment. Given that these adjustments were quite minor, we do not believe they have affected the results in any serious way. Second, to mimic the natural conditions for answering STs, we presented questions associated with each ST in blocks, because this is how each test would be answered in a natural context. Our results therefore do not clarify whether similar results would be obtained if questions associated with the four STs were randomly interspersed with one another during test-taking in the scanner. Third, future research using the present paradigm should consider including greater sample sizes to allow more robust estimates of effects. Although the sample size of 23 was sufficient to show reasonable, test-related regions, much larger samples would reveal other neural activity, allow other analysis techniques, allow subgroups differing in diagnosis or other individual differences to be compared, and in general would produce more definitive results. More specifically, we did not have enough participants with and without a diagnosis of PTSD for statistical comparisons; this is now generally accepted to be at least 20 per group for comparisons that are not a result of random assignment. In addition, increased sample sizes and the inclusion of populations with different symptoms and diagnoses would allow correlations among neural activity and individual differences measures of the participants

taken outside of the imaging environment to be investigated more easily. Fourth, because brain imaging during STs is a new paradigm introduced here, we were limited in the conclusions we could draw by a lack of studies of STs taken during neuroimaging from which specific activity could be predicted. A future literature of brain imaging studies of the processes involved in answering STs would extend the present work allowing specific predictions and triangulation on the processes underlying STs. Although a lack of prior research limited specific anatomical predictions, our study is an initial demonstration for lines of research developing theoretically and practically important test-taking neuroimaging paradigms.

FUTURE DIRECTIONS

There is a wealth of possible extensions to this research; we list a few examples. First, our study is an example of bridging levels of analysis related to mental health. In an attempt to understand psychopathology in terms of continuous differences in underlying processes as opposed to distinct diagnostic categories, the Research Domain Criteria (RDoC) system of integrating of levels of analysis has been established. The levels are genes, molecules, cells, circuits (the fMRI level we have been studying), physiology, behavior, self-reports (including STs), and paradigms (https://www.nimh.nih.gov/research-priorities/rdoc/constructs/rdoc-matrix.shtml). Thus, both this study and those we see as future studies are a direct analysis of self-report measures using the circuits level. Systematic examination of STs in a large homogenous sample that included participants with a range of disorders and comorbidities seems to be one obvious way to advance the goal of bridging levels of analysis in mental health.

Second, one important finding from this study was that the default and mentalization networks were associated with answering STs related to PTSD, that is, the PCL, CES, and Attachment. Future research should clarify the extent to which this finding extends beyond these three tests. For example, will similar activations of the default and mentalizing networks be seen in relation to answering other STs involving retrospective assessments of one's own behavior and mental states, such as STs addressing symptoms of depression, autobiographical memory, and mood states, or are these findings specific to the tests included in this study?

Third, related to the previous point, the findings stress that taking our mSTs activates area common in a subset of ToM tasks. Many of these tasks are related to deficits in ToM in PTSD including empathetic responding, recognizing facial emotions, and answering questions about social relations that were implicit in videos; deficits that in the context of family situations have been related to interpersonal violence, marital functioning, and intergenerational transmission of traumas [Lanius et al., 2014]. Thus, although we cannot claim that the activity in taking tests is directly related to having PTSD, the results draw attention to ToM deficits that could serve to increase the frequency of traumas and a decrease in social support that would help to alleviate symptoms. Similar deficits of ToM in major depressive disorder [Bora and Berk, 2016] and anxiety disorders [Washburn et al., 2016], which are often comorbid with PTSD, could have similar effects.

Fourth, studies varying both populations (e.g., ages, gender, clinical diagnosis, extremes of ability on the STs being investigated), experimental manipulations (e.g., pharmacological, mood change, sleep deprivation, and provocation studies), and manipulations of the test material would allow inferences we could not make. For example, we know that adolescents and older adults differ in the influences of their social groups and in risk-taking in ways that are relevant to the onset of addictions. Examining STs that measure these concepts while varying the age of the participant, either continuously or between groups, might provide insights into the development of the cognitive and emotional neural processes involved. However, it is also possible to include experimental manipulations that randomly assign participants to conditions, in which their social group or risky behavior is primed by the materials presented before or even while the mSTs are being taken. Combining a range of populations with experimental manipulations such as these is standard for many fMRI studies. For STs, it would open a way to answer questions about underlying neural processes in test-taking in a novel and efficient manner.

Fifth, we have concentrated on paper-and-pencil type standardized tests. However, this could be extended to structured interviews, including interviews done by a trained clinician, such as the CAPS used here to diagnose PTSD. For the fMRI environment, head movement is an issue, but no verbal responses are needed from the participant. Timing would be to the words spoken by the interviewer and manual responses of the participant to indicate a response or that they are finished thinking about what the interviewer said. Just as little is known about the neural processes of taking of STs, little is known about the neural processes that underlie interviews. Given the findings of social areas being involved in the mSTs used here, the effects of having an actual person met before entering the scanner, and of systematic variation of the age, gender, race, ethnicity, professional status, or other aspects the interviewer and how they match to the participant would be especially relevant.

Sixth, given the longstanding and politically and practically charged controversies over the meaning of test scores in education and job selection, we could add another level of analysis by examining whether participants with extreme scores recruit different patterns of neural activity. Observed differences in the neural processes could provide insights into test-taking behavior and skills. Online feedback using fMRI or electrical recordings might be able to change these. For STs that measure abilities and achievements, we often have additional knowledge about their neural substrates because the items are tasks that are

often studied independent of their role in STs. Verbal and subject-specific STs used in college and postgraduate admissions, especially those that measure vocabulary, are examples. For such ability and achievement tests, the added knowledge from the task itself should make many of the above manipulations easier to interpret. For instance, what we know about the neural basis of word meaning from neuroimaging and neuropsychology could be used to understand differences in the neural activity of verbal ability and achievement STs. Variables not thought to be specific to the neural substrates of the ability and achievement STs being tested, such as those related to sleep deprivation or stereotype threat, could be manipulated to ask if neural activity decreases in task-relevant areas and increases in areas related to the manipulation. Because ability and achievement tests are among those most often involved in current political debates about testing, providing such neural- and behavioral-level evidence would be timely and important.

Seventh, the methods used here also have the promise of assessing when specific STs are inappropriate for specific populations. For instance, it is well known that mentalization abilities are reduced in borderline personality disorder and in autistic populations. To the extent that answering specific psychometric tests recruits brain areas associated with mentalization, this might suggest that such tests may be less suitable for such populations or need modification. Here the neuroimaging would suggest investigations of causal relationships as the next step to address these issues. It may even be possible to inform when specific STs are inappropriate for specific populations based on the neural overlap of regions needed for the STs and areas that are damaged or that fail to function fully in a population. Decisions about the appropriateness of STs made on such neural bases would require much more neuroimaging evidence including neurally based norms established in an appropriate reference population. For any practical use, the efficacy of the neural findings for diagnosis and treatment would have to exceed that of behavioral measures. Nonetheless, this remains a potential goal for future research. In addition to the deficits in mentalization, other possible examples include concussions, early childhood deprivation leading to insecure attachment or general cognitive loss, and neuropsychological diagnoses. In all cases, the threshold for accepting evidence that affects treatment is much higher than and involves more ethical issues than research aimed at scientific understanding; it is a much more ambitious and distant goal.

CONCLUSION

We provide what we believe is the first systematic attempt to study the neural activity underlying the processes used in taking STs, tests which are carefully developed and tested behaviorally and which affect individuals' lives in important ways. The results provide an initial demonstration that STs can be studied productively using the techniques of human brain mapping. In our study, a pattern classification analysis classified correctly which questionnaire a participant was taking based on other participants' neural activity. The correlation between the distribution of classification errors and the semantic similarity of the test items indicated a high degree of similarity, demonstrating that the predictive information contained in patterns of neural activity was largely associated with the semantics of the tests, as opposed to unrelated factors that could be opportunistically leveraged in the pattern classification. Comparisons of our four mSTs to a General Knowledge baseline yielded systematic results that could be interpreted in terms of existing meta-analyses of neural activity in similar tasks and activated regions.

Turning to the content of our findings, an important result was that nearly all the activity in our mSTs coincided with areas found in meta-analyses of ToM. The activity in the PCL, CES, and Attachment compared to General Knowledge were different from each other, but all had most of their activity in regions that overlapped with the mentalizing network, and only in a fairly focused subset of the network that had obvious relevance to the particular mST. This novel finding stresses the mentalizing nature of included tests. A second finding was that the neural correlates of the cognitive and affective processes involved in completing our mSTs were independent of the scores they produced and of the participant's PTSD symptom severity. Thus, people who scored at different levels on the PTSD-related mSTs or who had different degrees of symptom severity did not vary on the neural processes they used. Both findings provide novel, theoretically and practically important information.

Our results point to many other areas where human brain mapping could make significant contributions. By including both a range of participant populations and experimental manipulations, standard procedures for fMRI research that we did not use in this first demonstration study, many questions become tractable. Applications are possible in probing the neural processes involved in STs related to academic and professional knowledge and abilities, personality and other traits, opinions and attitudes, health, and other areas. Applications are also possible to study how the neural processes are affected by individual differences factors related to the people taking the test, as well as their training and test-taking conditions. Moreover, uncovering differences in neural processes associated with different populations and different conditions would render it possible to decide whether the test results across such factors are measuring similar neural processes and can be fairly compared to each other.

AUTHOR CONTRIBUTIONS

All authors contributed to the study design, interpretation of the results, the conceptual development of the

project throughout its progress, and commented on the later drafts. D. C. Rubin developed the basic idea for the study and was responsible for drafting the nontechnical sections and integrating the research and article. S. A. Hall was in charge of the behavioral and fMRI data collection, and the quality control of the fMRI data. D. Li, P. A. Kragel, and S. A. Hall contributed to the data analysis and consulted with each other as needed. P. A. Kragel performed the data analysis, drafted the manuscript sections, and constructed the figures for the separability and Neurosynth analyses. D. Li performed the data analysis, drafted the manuscript sections, and constructed the figures for the rest of the article. All authors approved the final version of the manuscript for submission.

ACKNOWLEDGMENT

The authors thank Tobias Egner, Christin Ogle, and Kaitlyn Brodar for comments, and Kaitlyn Brodar's help in running the study.

CONFLICT OF INTERESTS

We have no conflicts of interest.

REFERENCES

Andrews-Hanna JR, Saxe R, Yarkoni T (2014): Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. NeuroImage 91:324–335.

Apps MA, Lockwood PL, Balsters JH (2013): The role of the midcingulate cortex in monitoring others' decisions. Front Neurosci 7:251.

Association AP (2013): Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub.

Berntsen D, Rubin DC (2006): The centrality of event scale: A measure of integrating a trauma into one's identity and its relation to post-traumatic stress disorder symptoms. Behav Res Therapy 44:219–231.

Berntsen D, Rubin DC (2007): When a trauma becomes a key to identity: Enhanced integration of trauma memories predicts posttraumatic stress disorder symptoms. Appl Cogn Psychol 21:417–431.

Blake DD, Weathers FW, Nagy LM, Kaloupek DG, Gusman FD, Charney DS, Keane TM (1995): The development of a Clinician-Administered PTSD Scale. J Trauma Stress 8:75–90.

Bora E, Berk M (2016): Theory of mind in major depressive disorder: A meta-analysis. J Affect Disord 191:49–55.

Bowlby J (1969): Attachment and Loss: Attachment; John Bowlby. Basic Books.

Brennan KA, Clark CL, Shaver PR (1998): Self-report measurement of adult attachment: An integrative overview.

Brewer JA, Garrison KA, Whitfield-Gabrieli S (2013): What about the "self" is processed in the posterior cingulate cortex?

Brown VM, LaBar KS, Haswell CC, Gold AL, Beall SK, Van Voorhees E, Marx CE, Calhoun PS, Fairbank JA, Green KT (2014): Altered resting-state functional connectivity of basolateral and centromedial amygdala complexes in posttraumatic stress disorder. Neuropsychopharmacology 39:361–369.

Buckner RL, Andrews-Hanna JR, Schacter DL (2008): The brain's default network: Anatomy, function, and relevance to disease. Ann N Y Acad Sci 1124:1–38.

Buckner RL, Carroll DC (2007): Self-projection and the brain. Trends Cogn Sci 11:49–57.

Bzdok D, Schilbach L, Vogeley K, Schneider K, Laird AR, Langner R, Eickhoff SB (2012): Parsing the neural correlates of moral cognition: ALE meta-analysis on morality, theory of mind, and empathy. Brain Struct Funct 217:783–796.

Carter RM, Huettel SA (2013): A nexus model of the temporal-parietal junction. Trends Cogn Sci 17:328–336.

Charuvastra A, Cloitre M (2008): Social bonds and posttraumatic stress disorder. Annu Rev Psychol 59:301–328.

Costa PT, McCrae RR (2008): The revised NEO personality inventory (NEO-PI-R). In: Boyle GJ, Matthews G, Saklofske DH, editors. The SAGE Handbook of Personality Theory and Assessment London: Sage Publications Ltd, 2:179–198.

Denny BT, Kober H, Wager TD, Ochsner KN (2012): A meta-analysis of functional neuroimaging studies of self-and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. J Cogn Neurosci 24:1742–1752.

Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995): Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. Magn Reson Med 33:636–647.

Garrett AS, Carrion V, Kletter H, Karchemskiy A, Weems CF, Reiss A (2012): Brain activation to facial expressions in youth with PTSD symptoms. Depression Anxiety 29:449–459.

Griffin DW, Bartholomew K (1994): The metaphysics of measurement: The case of adult attachment.

Hall SA, Rubin DC, Miles A, Davis SW, Wing EA, Cabeza R, Berntsen D (2014): The neural basis of involuntary episodic memories. J Cogn Neurosci 26:2385–2399.

Han L, Kashyap A, Finin T, Mayfield J, Weese J (2013): UMBC EBIQUITY-CORE: Semantic textual similarity systems. pp 44–52.

Hennig-Fast K, Werner NS, Lermer R, Latscha K, Meister F, Reiser M, Engel RR, Meindl T (2009): After facing traumatic stress: Brain activation, cognition and stress coping in policemen. J Psychiatr Res 43:1146–1155.

John OP, Donahue EM, Kentle RL (1991): The Big Five Inventory—Versions 4a and 54. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.

Kilpatrick DG, Resnick HS, Milanak ME, Miller MW, Keyes KM, Friedman MJ (2013): National estimates of exposure to traumatic events and PTSD prevalence using DSM-IV and DSM-5 criteria. J Trauma Stress 26:537–547.

Kim H (2012): A dual-subsystem model of the brain's default network: Self-referential processing, memory retrieval processes, and autobiographical memory retrieval. NeuroImage 61: 966–977.

Kohavi R (1995): A study of cross-validation and bootstrap for accuracy estimation and model selection. In Stanford, CA. pp 1137–1145.

Lancaster JL, Tordesillas-Gutiérrez D, Martinez M, Salinas F, Evans A, Zilles K, Mazziotta JC, Fox PT (2007): Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum Brain Mapp 28:1194–1205.

Lanius R, Frewen P, Nazarov A, McKinnon M (2014): A social-cognitive–neuroscience approach to PTSD: Clinical and research perspectives. Neurobiol Treat Trauma Dissoc 69–80.

Mar RA (2011): The neural bases of social cognition and story comprehension. Annu Rev Psychol 62:103–134.

Mikulincer M, Shaver PR (2007): Attachment in Adulthood: Structure, Dynamics, and Change. Guilford Press.

Ogle CM, Rubin DC, Siegler IC (2015): The relation between insecure attachment and posttraumatic stress: Early life versus adulthood traumas. Psychol Trauma Theory Res Pract Policy 7:324.

Olson IR, McCoy D, Klobusicky E, Ross LA (2013): Social cognition and the anterior temporal lobes: A review and theoretical framework. Soc Cogn Affect Neurosci 8:123–133.

Olson IR, Plotzker A, Ezzyat Y (2007): The enigmatic temporal pole: A review of findings on social and emotional processing. Brain 130:1718–1731.

Palomero-Gallagher N, Vogt BA, Schleicher A, Mayberg HS, Zilles K (2009): Receptor architecture of human cingulate cortex: Evaluation of the four-region neurobiological model. Hum Brain Mapp 30:2336–2355.

Raichle ME (2015): The brain's default mode network. Annu Rev Neurosci 38:433–447.

Rice GE, Ralph MAL, Hoffman P (2015): The roles of left versus right anterior temporal lobes in conceptual knowledge: An ALE meta-analysis of 97 functional neuroimaging studies. Cereb Cortex bhv024.

Roberts AL, Gilman SE, Breslau J, Breslau N, Koenen KC (2011): Race/ethnic differences in exposure to traumatic events, development of post-traumatic stress disorder, and treatment-seeking for post-traumatic stress disorder in the United States. Psychol Med 41:71–83.

Rubin DC, Berntsen D, Bohni MK (2008): A memory-based model of posttraumatic stress disorder: Evaluating basic assumptions underlying the PTSD diagnosis. Psychol Rev 115:985.

Rubin DC, Boals A, Hoyle RH (2014): Narrative centrality and negative affectivity: Independent and interactive contributors to stress reactions. J Exp Psychol Gen 143:1159.

Schaafsma SM, Pfaff DW, Spunt RP, Adolphs R (2015): Deconstructing and reconstructing theory of mind. Trends Cogn Sci 19:65–72.

Schwartz SJ, Lilienfeld SO, Meca A, Sauvigné KC (2016): The role of neuroscience within psychology: A call for inclusiveness over exclusiveness. Am Psychol 71:52.

Seghier ML (2013): The angular gyrus multiple functions and multiple subdivisions. The. Neuroscientist 19:43–61.

Shin LM, Whalen PJ, Pitman RK, Bush G, Macklin ML, Lasko NB, Orr SP, McInerney SC, Rauch SL (2001): An fMRI study of anterior cingulate function in posttraumatic stress disorder. Biol Psychiat 50:932–942.

Shirer W, Ryali S, Rykhlevskaia E, Menon V, Greicius M (2012): Decoding subject-driven cognitive states with whole-brain connectivity patterns. Cereb Cortex 22:158–165.

Song X-W, Dong Z-Y, Long X-Y, Li S-F, Zuo X-N, Zhu C-Z, He Y, Yan C-G, Zang Y-F (2011): REST: A toolkit for resting-state functional magnetic resonance imaging data processing. PLoS One 6:e25031.

Sperduti M, Delaveau P, Fossati P, Nadel J (2011): Different brain structures related to self-and external-agency attribution: A brief review and meta-analysis. Brain Struct Funct 216:151–157.

Spitzer RL, Kroenke K, Williams JB, Löwe B (2006): A brief measure for assessing generalized anxiety disorder: The GAD-7. Arch Int Med 166:1092–1097.

Spreng RN, Mar RA, Kim AS (2009): The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: A quantitative meta-analysis. J Cogn Neurosci 21:489–510.

St. Jacques PLS, Botzung A, Miles A, Rubin DC (2011): Functional neuroimaging of emotionally intense autobiographical memories in post-traumatic stress disorder. J Psychiatr Res 45:630–637.

St. Jacques PLS, Kragel PA, Rubin DC (2013): Neural networks supporting autobiographical memory retrieval in posttraumatic stress disorder. Cogn Affect Behav Neurosci 13:554–566.

Tauber SK, Dunlosky J, Rawson KA, Rhodes MG, Sitzman DM (2013): General knowledge norms: Updated and expanded from the Nelson and Narens (1980) norms. Behav Res Methods 45:1115–1143.

Van Overwalle F (2009): Social cognition and the brain: A meta-analysis. Hum Brain Mapp 30:829–858.

Vogt BA, Nimchinsky EA, Vogt LJ, Hof PR (1995): Human cingulate cortex: Surface features, flat maps, and cytoarchitecture. J Comp Neurol 359:490–506.

Washburn D, Wilson G, Roes M, Rnic K, Harkness KL (2016): Theory of mind in social anxiety disorder, depression, and comorbid conditions. J Anxiety Disord 37:71–77.

Weathers FW, Litz BT, TM K, Palmieri PA, Marx BP, Schnurr PP (2013): The PTSD Checklist for DSM-5 (PCL-5). Scale available from the National Center for PTSD at www.ptsd.va.gov.

Wong C, Gallate J (2012): The function of the anterior temporal lobe: A review of the empirical evidence. Brain Res 1449:94–116.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011): Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8:665–670.

Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zollei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011): The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106:1125–1165.